

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/525,020A
Source: IFW/6
Date Processed by STIC: 7-14-06

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 07/14/2006

PATENT APPLICATION: US/10/525,020A

TIME: 09:30:49

Input Set : A:\81356233.APP

Output Set: N:\CRF4\07142006\J525020A.raw

3 <110> APPLICANT: YAMAMOTO, KAZUO
 4 SATO, AKIRA
 5 SHIMAUCHI, JUNKO
 6 MATSUMOTO, MARIKO
 8 <120> TITLE OF INVENTION: CARBOHYDRATE LIBRARY CONSTRUCTED BY GENE ALTERATION OF
 9 CARGO RECEPTORS
 11 <130> FILE REFERENCE: 081356-0233
 13 <140> CURRENT APPLICATION NUMBER: 10/525,020A
 14 <141> CURRENT FILING DATE: 2005-02-18
 16 <150> PRIOR APPLICATION NUMBER: PCT/JP03/01718
 17 <151> PRIOR FILING DATE: 2003-02-18
 19 <150> PRIOR APPLICATION NUMBER: JP 2002-238559
 20 <151> PRIOR FILING DATE: 2002-08-19
 22 <160> NUMBER OF SEQ ID NOS: 33
 24 <170> SOFTWARE: PatentIn Ver. 3.3
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 2768
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Homo sapiens
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (22)..(1554)
 35 <400> SEQUENCE: 1
 36 ggtcgcggttc cagaatccaa g atg gcg gga tcc agg caa agg ggt ctc cgg 51
 37 Met Ala Gly Ser Arg Gln Arg Gly Leu Arg
 38 1 5 10
 40 gcc aga gtt cgg ccg ctg ttc tgc gcc ttg ctg ctg tca ctc ggt cgc 99
 41 Ala Arg Val Arg Pro Leu Phe Cys Ala Leu Leu Leu Ser Leu Gly Arg
 42 15 20 25
 44 ttc gtc cgg ggc gac ggc gtg gga gga gac ccc gcg gtc gcg ttg cca 147
 45 Phe Val Arg Gly Asp Gly Val Gly Gly Asp Pro Ala Val Ala Leu Pro
 46 30 35 40
 48 cat cgc cgt ttc gag tac aaa tac agc ttc aag ggg ccg cac ctg gtg 195
 49 His Arg Arg Phe Glu Tyr Lys Tyr Ser Phe Lys Gly Pro His Leu Val
 50 45 50 55
 52 cag agc gac ggg acc gtg ccc ttc tgg gcc cac gcg ggg aat gct att 243
 53 Gln Ser Asp Gly Thr Val Pro Phe Trp Ala His Ala Gly Asn Ala Ile
 54 60 65 70
 56 cca agt tca gat caa att cga gta gca cca tct tta aaa agc caa aga 291
 57 Pro Ser Ser Asp Gln Ile Arg Val Ala Pro Ser Leu Lys Ser Gln Arg
 58 75 80 85 90
 60 ggc tca gtg tgg aca aag aca aaa gcg gcc ttt gag aac tgg gaa gtt 339
 61 Gly Ser Val Trp Thr Lys Thr Lys Ala Ala Phe Glu Asn Trp Glu Val

(pg.6)

RAW SEQUENCE LISTING

DATE: 07/14/2006

PATENT APPLICATION: US/10/525,020A

TIME: 09:30:50

Input Set : A:\81356233.APP

Output Set: N:\CRF4\07142006\J525020A.raw

62		95		100		105	
64	gag	gtg	aca	ttt	cga	gtg	act
65	Glu	Val	Thr	Phe	Arg	Val	Thr
66				110			115
68	cta	gca	att	tgg	tat	gca	gaa
69	Leu	Ala	Ile	Trp	Tyr	Ala	Glu
70				125			130
72	gga	tca	gct	gat	ctg	tgg	aat
73	Gly	Ser	Ala	Asp	Leu	Trp	Asn
74				140			145
76	gac	aat	gat	gga	aag	aaa	aat
77	Asp	Asn	Asp	Gly	Lys	Lys	Asn
78	155						160
80	aat	gga	caa	atc	cat	tat	gac
81	Asn	Gly	Gln	Ile	His	Tyr	Asp
82				175			180
84	ttg	gca	agt	tgc	cag	agg	gac
85	Leu	Ala	Ser	Cys	Gln	Arg	Asp
86				190			195
88	gca	aag	att	acc	tat	tac	cag
89	Ala	Lys	Ile	Thr	Tyr	Tyr	Gln
90				205			210
92	ggc	ttt	aca	cca	gat	aaa	aat
93	Gly	Phe	Thr	Pro	Asp	Lys	Asn
94				220			225
96	aat	atg	att	atc	cct	gca	caa
97	Asn	Met	Ile	Ile	Pro	Ala	Gln
98	235						240
100	gga	ggt	ctt	gca	gat	gac	cat
101	Gly	Gly	Leu	Ala	Asp	Asp	His
102				255			260
104	ttg	act	gaa	cct	gga	aaa	gag
105	Leu	Thr	Glu	Pro	Gly	Lys	Glu
106				270			275
108	gaa	aag	gaa	aaa	gaa	aag	tat
109	Glu	Lys	Glu	Lys	Glu	Lys	Tyr
110				285			290
112	gaa	ttg	gat	aaa	aaa	aaa	gag
113	Glu	Leu	Asp	Lys	Lys	Lys	Glu
114				300			305
116	caa	ggg	cag	cct	gcg	gag	gaa
117	Gln	Gly	Gln	Pro	Ala	Glu	Glu
118	315						320
120	cta	aga	caa	gtc	ttt	gaa	gga
121	Leu	Arg	Gln	Val	Phe	Glu	Gly
122				335			340
124	cag	ctg	aac	cgg	cag	tta	gat
125	Gln	Leu	Asn	Arg	Gln	Leu	Asp
126				350			355

RAW SEQUENCE LISTING

DATE: 07/14/2006

PATENT APPLICATION: US/10/525,020A

TIME: 09:30:50

Input Set : A:\81356233.APP

Output Set: N:\CRF4\07142006\J525020A.raw

```

128 gtc tct tcc tta aca gag gaa atc tct aaa aga gga gca gga atg cct 1155
129 Val Ser Ser Leu Thr Glu Glu Ile Ser Lys Arg Gly Ala Gly Met Pro
130      365      370      375
132 ggg cag cat ggg cag att act caa caa gaa ctg gat act gtt gtg aaa 1203
133 Gly Gln His Gly Gln Ile Thr Gln Gln Glu Leu Asp Thr Val Val Lys
134      380      385      390
136 act cag cat gag att ctg aga caa gta aat gaa atg aaa aat tcc atg 1251
137 Thr Gln His Glu Ile Leu Arg Gln Val Asn Glu Met Lys Asn Ser Met
138 395      400      405      410
140 agt gaa acc gtc aga ctg gtc agt gga atg cag cac cct ggc tct gct 1299
141 Ser Glu Thr Val Arg Leu Val Ser Gly Met Gln His Pro Gly Ser Ala
142      415      420      425
144 gga ggc gtc tat gag aca aca cag cac ttc att gac atc aaa gag cac 1347
145 Gly Gly Val Tyr Glu Thr Thr Gln His Phe Ile Asp Ile Lys Glu His
146      430      435      440
148 ctg cac ata gta aag agg gac ata gat aac tta gtg cag cga aat atg 1395
149 Leu His Ile Val Lys Arg Asp Ile Asp Asn Leu Val Gln Arg Asn Met
150      445      450      455
152 cca tca aat gaa aag ccg aaa tgc cca gaa cta cca cca ttt cca tca 1443
153 Pro Ser Asn Glu Lys Pro Lys Cys Pro Glu Leu Pro Pro Phe Pro Ser
154      460      465      470
156 tgt ttg tct acg gtc cac ttc att ata ttt gtt gtg gtg caa act gta 1491
157 Cys Leu Ser Thr Val His Phe Ile Ile Phe Val Val Val Gln Thr Val
158 475      480      485      490
160 tta ttc att ggt tat atc atg tat agg tct cag caa gaa gca gct gcc 1539
161 Leu Phe Ile Gly Tyr Ile Met Tyr Arg Ser Gln Gln Glu Ala Ala Ala
162      495      500      505
164 aaa aaa ttc ttt tga ctaccatttt cctgtgtact tcattctatatt gtgtacaaaa 1594
165 Lys Lys Phe Phe
166      510
168 tgagtcgttt tgagggaatt taagtattta aattgcttca tagtctaaat tattaatttt 1654
170 ctttaataaaa taactgttta aacattgatt tgcagttaag aataaacctt aaagcaaaga 1714
172 caaccacatt ttaatttggt cacagtatgt aaatctgtct aaatttcagt gaatttctgg 1774
174 tcagtatgat gcagcctctg agcagaatat tgaccagtaa gagggtaaat aaagtggggg 1834
176 caaccctgga tatgaatggt accccctaag tctccaatat tgcaggtttc cctgtataac 1894
178 gtaaacacac ttgccctcat gccctccaga atatgaggtc taattaagaa gtccatcagg 1954
180 tttattttgt aaccaaagtc ttttttagag gtcagacttc ctaatcaaag gccctgggct 2014
182 gcagtccttt tcattctaat gcaacttcct ttgaaatcaa agaataatttt gtctgagagc 2074
184 ttttaaggatc tggtaataga cttcaaaatg ttaagtgaag tttttttttc ctctatttat 2134
186 caatgatata tttcactttt aaaggaaatt ttagaggaaa attaatagct gctttttgca 2194
188 ctaaaaaacc ttgtgggtgg aaatattcct ctgagaatgg cttttatagg tattttgcct 2254
190 ggtaatgtat tcattcatga ttgcccatat tcttgaatgt ttcttcattc caatgggggtc 2314
192 aggtcaatat tatgaaaata atttttatat ttatatattgt aactaagaat ttattttctcc 2374
194 ctttactaca cgatgtaaat tcacgtcaaa ttcatgatgc tgaggattta aattcacaaa 2434
196 acctgccact acattctggt ttacattagt taactcatgc tggtctgggt tagtgaccat 2494
198 ttgcatactc ttttaaatca aggaggctgt agtagaggca gttttaagat tcttgaaggc 2554
200 aaaatttgaa aaacagtga tacttctaatt tgtttctttt tagtgccaga actaagacat 2614
202 tgtgaagcac ttgttagtaa acttaacctt gaaatgtcag actggaagga gtttttatgt 2674
204 ctttgtgcat acttctgggt attacagaaa cagtctgtaa ataacatttt aagatgcaaa 2734

```

RAW SEQUENCE LISTING

DATE: 07/14/2006

PATENT APPLICATION: US/10/525,020A

TIME: 09:30:50

Input Set : A:\81356233.APP

Output Set: N:\CRF4\07142006\J525020A.raw

206 ttttaattctg ttcacagctg atttatactg attt 2768

209 <210> SEQ ID NO: 2

210 <211> LENGTH: 510

211 <212> TYPE: PRT

212 <213> ORGANISM: Homo sapiens

214 <400> SEQUENCE: 2

215 Met Ala Gly Ser Arg Gln Arg Gly Leu Arg Ala Arg Val Arg Pro Leu

216 1 5 10 15

218 Phe Cys Ala Leu Leu Leu Ser Leu Gly Arg Phe Val Arg Gly Asp Gly

219 20 25 30

221 Val Gly Gly Asp Pro Ala Val Ala Leu Pro His Arg Arg Phe Glu Tyr

222 35 40 45

224 Lys Tyr Ser Phe Lys Gly Pro His Leu Val Gln Ser Ser Asp Gly Thr Val

225 50 55 60

227 Pro Phe Trp Ala His Ala Gly Asn Ala Ile Pro Ser Ser Asp Gln Ile

228 65 70 75 80

230 Arg Val Ala Pro Ser Leu Lys Ser Gln Arg Gly Ser Val Trp Thr Lys

231 85 90 95

233 Thr Lys Ala Ala Phe Glu Asn Trp Glu Val Glu Val Thr Phe Arg Val

234 100 105 110

236 Thr Gly Arg Gly Arg Ile Gly Ala Asp Gly Leu Ala Ile Trp Tyr Ala

237 115 120 125

239 Glu Asn Gln Gly Leu Glu Gly Pro Val Phe Gly Ser Ala Asp Leu Trp

240 130 135 140

242 Asn Gly Val Gly Ile Phe Phe Asp Thr Phe Asp Asn Asp Gly Lys Lys

243 145 150 155 160

245 Asn Asn Pro Ala Ile Val Ile Ile Gly Asn Asn Gly Gln Ile His Tyr

246 165 170 175

248 Asp His Gln Asn Asp Gly Ala Ser Gln Ala Leu Ala Ser Cys Gln Arg

249 180 185 190

251 Asp Phe Arg Asn Lys Pro Tyr Pro Val Arg Ala Lys Ile Thr Tyr Tyr

252 195 200 205

254 Gln Asn Thr Leu Thr Val Met Ile Asn Asn Gly Phe Thr Pro Asp Lys

255 210 215 220

257 Asn Asp Tyr Glu Phe Cys Ala Lys Val Glu Asn Met Ile Ile Pro Ala

258 225 230 235 240

260 Gln Gly His Phe Gly Ile Ser Ala Ala Thr Gly Gly Leu Ala Asp Asp

261 245 250 255

263 His Asp Val Leu Ser Phe Leu Thr Phe Gln Leu Thr Glu Pro Gly Lys

264 260 265 270

266 Glu Pro Pro Thr Pro Asp Lys Glu Ile Ser Glu Lys Glu Lys Glu Lys

267 275 280 285

269 Tyr Gln Glu Glu Phe Glu His Phe Gln Gln Glu Leu Asp Lys Lys Lys

270 290 295 300

272 Glu Glu Phe Gln Lys Gly His Pro Asp Leu Gln Gly Gln Pro Ala Glu

273 305 310 315 320

275 Glu Ile Phe Glu Ser Val Gly Asp Arg Glu Leu Arg Gln Val Phe Glu

276 325 330 335

278 Gly Gln Asn Arg Ile His Leu Glu Ile Lys Gln Leu Asn Arg Gln Leu

RAW SEQUENCE LISTING

DATE: 07/14/2006

PATENT APPLICATION: US/10/525,020A

TIME: 09:30:50

Input Set : A:\81356233.APP

Output Set: N:\CRF4\07142006\J525020A.raw

```

279          340          345          350
281 Asp Met Ile Leu Asp Glu Gln Arg Arg Tyr Val Ser Ser Leu Thr Glu
282          355          360          365
284 Glu Ile Ser Lys Arg Gly Ala Gly Met Pro Gly Gln His Gly Gln Ile
285          370          375          380
287 Thr Gln Gln Glu Leu Asp Thr Val Val Lys Thr Gln His Glu Ile Leu
288 385          390          395          400
290 Arg Gln Val Asn Glu Met Lys Asn Ser Met Ser Glu Thr Val Arg Leu
291          405          410          415
293 Val Ser Gly Met Gln His Pro Gly Ser Ala Gly Gly Val Tyr Glu Thr
294          420          425          430
296 Thr Gln His Phe Ile Asp Ile Lys Glu His Leu His Ile Val Lys Arg
297          435          440          445
299 Asp Ile Asp Asn Leu Val Gln Arg Asn Met Pro Ser Asn Glu Lys Pro
300          450          455          460
302 Lys Cys Pro Glu Leu Pro Pro Phe Pro Ser Cys Leu Ser Thr Val His
303 465          470          475          480
305 Phe Ile Ile Phe Val Val Val Gln Thr Val Leu Phe Ile Gly Tyr Ile
306          485          490          495
308 Met Tyr Arg Ser Gln Gln Glu Ala Ala Ala Lys Lys Phe Phe
309          500          505          510
313 <210> SEQ ID NO: 3
314 <211> LENGTH: 1407
315 <212> TYPE: DNA
316 <213> ORGANISM: Homo sapiens
318 <220> FEATURE:
319 <221> NAME/KEY: CDS
320 <222> LOCATION: (1)..(1071)
322 <400> SEQUENCE: 3
323 atg gcg gcg gaa ggc tgg att tgg cgt tgg ggc tgg ggc cgg cgg tgc 48
324 Met Ala Ala Glu Gly Trp Ile Trp Arg Trp Gly Trp Gly Arg Arg Cys
325 1 5 10 15
327 ctg gga agg cct ggg ctt ctc ggc ccc ggc cct ggc ccc act aca cct 96
328 Leu Gly Arg Pro Gly Leu Leu Gly Pro Gly Pro Gly Pro Thr Thr Pro
329 20 25 30
331 ctc ttt ctt ctt ttg ttg ttg ggg tct gtg act gcg gat ata act gac 144
332 Leu Phe Leu Leu Leu Leu Leu Gly Ser Val Thr Ala Asp Ile Thr Asp
333 35 40 45
335 ggc aac agt gaa cat ctc aag cgg gag cat tcg ctc att aag ccc tac 192
336 Gly Asn Ser Glu His Leu Lys Arg Glu His Ser Leu Ile Lys Pro Tyr
337 50 55 60
339 caa ggg gtc ggt tcc agc tct atg ccc ctc tgg gac ttc cag ggc agc 240
340 Gln Gly Val Gly Ser Ser Ser Met Pro Leu Trp Asp Phe Gln Gly Ser
341 65 70 75 80
343 act atg ctc acg agc cag tac gta cgt ctg acc cct gac gag cgc agc 288
344 Thr Met Leu Thr Ser Gln Tyr Val Arg Leu Thr Pro Asp Glu Arg Ser
345 85 90 95
347 aaa gag ggc tct atc tgg aac cac cag ccg tgc ttc ctc aaa gac tgg 336
348 Lys Glu Gly Ser Ile Trp Asn His Gln Pro Cys Phe Leu Lys Asp Trp

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/525,020A

DATE: 07/14/2006
TIME: 09:30:51

Input Set : A:\81356233.APP
Output Set: N:\CRF4\07142006\J525020A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 13, 14, 16, 17, 18, 20, 25, 26, 28, 29, 31, 32, 34, 35

Seq#:14; N Pos. 12, 13, 15, 16, 18, 19, 24, 25, 27, 28, 30, 31, 33, 34

VERIFICATION SUMMARY

DATE: 07/14/2006

PATENT APPLICATION: US/10/525,020A

TIME: 09:30:51

Input Set : A:\81356233.APP

Output Set: N:\CRF4\07142006\J525020A.raw

L:574 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0

L:700 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0